

Семинар «Прикладная статистика» (ИМ СО РАН)

15 апреля 2024 года

**Экономические алгоритмы
моделирования одномерных непрерывных
случайных величин, основанные
на принципе «уравнивания вероятностей»**

Войтишек Антон Вацлавович
(ИВМиМГ СО РАН, г. Новосибирск)

Брызгалов Виктор Леонидович
(Лицей № 130, г. Новосибирск)

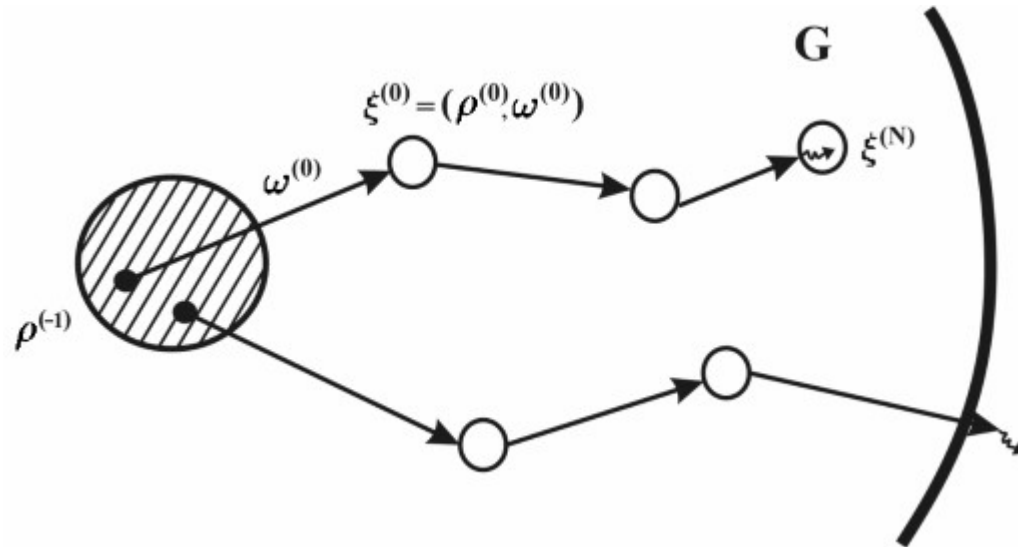
ПЛАН ДОКЛАДА

1. О «необычном» стиле доклада (для семинара «Прикладная статистика»).
2. О необходимости получения многочисленных, смоделированных на компьютере выборочных значений одномерных случайных величин при реализации основной схемы метода Монте-Карло для решения практически значимых задач.
3. Метод обратной функции распределения: преимущества и ограничения.
4. Двусторонний метод исключения как альтернатива методу обратной функции распределения.
5. Обоснование мажорантного метода исключения. Двусторонний метод.
6. Использование кусочно-постоянных мажорант и минорант (для монотонных плотностей на конечных интервалах распределения).
7. Моделирование вспомогательной случайной величины по кусочно-постоянной плотности, пропорциональной мажоранте. Использование теории модифицированного метода дискретной суперпозиции и уравнивания вероятностей.
8. Компьютерная система EDSRM (Economical Double-Sided Rejection Method см. <http://edsrm.nikita-e.ru/>). Примеры сравнения экономичного двустороннего метода и метода обратной функции распределения (степенное распределение и др.).

2. О необходимости получения многочисленных, смоделированных на компьютере выборочных значений одномерных случайных величин при реализации основной схемы метода Монте-Карло для решения практически значимых задач.

Широкое распространение получают **стохастические (вероятностные) модели**. Им соответствуют вычислительные (компьютерные) алгоритмы статистического моделирования (или **методы Монте-Карло**); здесь и далее приведены сведения из книги [1].

Рассмотрим классический пример «практически значимой» задачи: **модель переноса излучения**.



Модель переноса излучения

Основная схема метода Монте-Карло (ММК) выглядит следующим образом.

Пусть требуется приближенное вычисление величины I .

Предполагается возможность выбора такой случайной величины (**монте-карловской оценки, оценителя**) ζ , для которой $I = \mathbf{E}\zeta$ и выборочные значения $\zeta_1, \dots, \zeta_n$ случайной величины ζ могут быть достаточно эффективно (экономично) получены на компьютере. Используя закон больших чисел, строим приближение

$$I = \mathbf{E}\zeta \approx Z_n = \frac{\zeta_1 + \dots + \zeta_n}{n}.$$

Средняя погрешность ММК

$$\delta_n = \mathbf{E}|I - Z_n| \approx H \frac{\sqrt{\mathbf{D}\zeta}}{\sqrt{n}}; \quad H = \sqrt{\frac{2}{\pi}} \approx 0,80 \dots$$

(НЕДАВНЕЕ «ОТКРЫТИЕ»); здесь $\mathbf{D}\zeta = \mathbf{E}|\zeta - \mathbf{E}\zeta|^2$ – дисперсия случайной величины ζ (при выводе этого соотношения используется центральная предельная теорема).

Скорость сходимости основной схемы ММК при $n \rightarrow \infty$ невелика, но «стабильна» (например, не зависит от размерности задачи).

В приведенной выше основной схеме ММК

$$I = \mathbf{E}\zeta \approx Z_n = \frac{\zeta_1 + \dots + \zeta_n}{n}$$

случайная величина (монте-карловская оценка, оценщик) ζ имеет вид $\zeta = q(\xi)$ (например, при приближении интеграла мы имели $q(\xi) = \frac{g(\xi)}{f_\xi(\xi)}$; $\xi \sim f_\xi(\mathbf{x})$).

В свою очередь, для получения значений $\zeta_i = q(\xi_i)$ нужно уметь получать на компьютере выборочные значения случайного вектора $\xi = (\xi^{(1)}, \dots, \xi^{(d)})$ согласно заданной совместной плотности $f_\xi(\mathbf{x})$.

Дело сводится к моделированию компонент $\xi^{(1)}, \dots, \xi^{(d)}$ (одномерных случайных величин) согласно условным плотностям из разложения плотности $f_\xi(\mathbf{x})$.

Учитывая соотношение $\delta_n \approx H \frac{\sqrt{\mathbf{D}\zeta}}{\sqrt{n}}$, весьма желательно, чтобы описанное моделирование случайных компонент $\xi^{(1)}, \dots, \xi^{(d)}$ было экономичным.

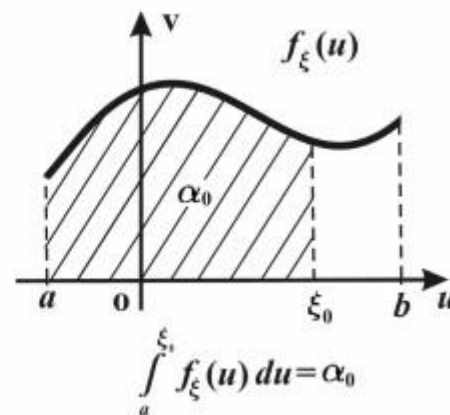
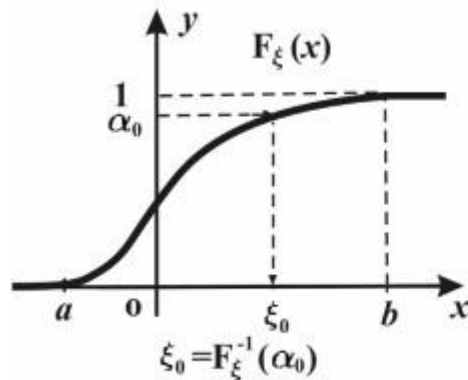
3. Метод обратной функции распределения: преимущества и ограничения.

Исторически основным методом моделирования одномерной случайной $\xi \in (a, b)$ величины согласно заданной плотности $f_\xi(u); u \in (a, b)$ является **формула метода обратной функции распределения**

$$\xi_0 = F_\xi^{-1}(\alpha_0); \quad F_\xi(x) = \mathbf{P}\{\xi < x\}, \quad \alpha_0 \in U(0,1), \quad (*)$$

получающейся из решения уравнения $\int_a^{\xi_0} f_\xi(u) du = \alpha_0$ (**).

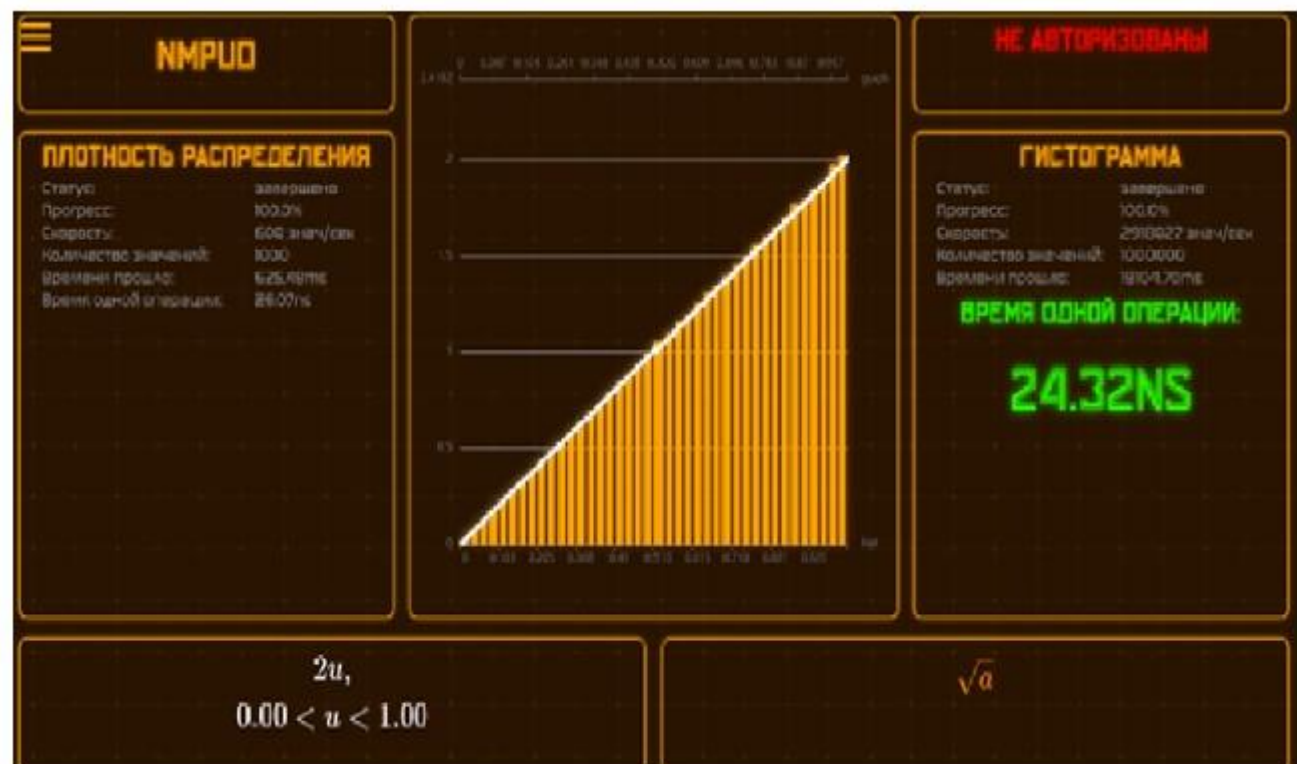
В формулах (*) и (**) $\alpha_0 \in U(0,1)$ – стандартное случайное число (выборочное значение случайной величины α , равномерно распределенной на интервале (0,1) – получается с помощью специальных подпрограмм; в Новосибирске – метод вычетов).



Далеко не для всех плотностей $f_{\xi}(u)$ уравнение $\int_a^{\xi_0} f_{\xi}(u) du = \alpha_0$ разрешимо в элементарных функциях. Определенным «методическим открытием» на курсах по ММК явилась **технология последовательных (вложенных) замен**, основанная на формулах

$$f_{\xi}(u) = f_{\eta}[\varphi(u)]|\varphi'(u)|; u \in (a, b); \varphi(u): (c, d) \rightarrow (a, b);$$
$$\xi_0 = \varphi^{-1}[\Psi_{\eta}(\alpha_0)].$$

Ограничения по применению этой технологии удалось изучить с помощью системы **NMPUD (Numerical Modelling of Probabilistic Univariate Distributions** – автор Черкашин Данил Андреевич, студент второго курса ВМК Московского государственного университета) – ссылка <https://nmpud.netlify.app>.



Что удалось выяснить с помощью системы NMPUD – см. работу

Войтишек А. В., Гаджихмедов М. Г., Рыжов И. А., Трофимов И. А. Сравнение вычислительных затрат на основные математические операции с помощью системы NMPUD // Информационные технологии и математическое моделирование (ИТММ-2021): Материалы XX Международной конференции имени А. Ф. Терпугова (1–5 декабря 2021 года). – Томск: Издательство Томского государственного университета, 2022. – С. 329 – 334 :

- компьютерные затраты на сложение, умножение и деление примерно одинаковы (часто даже умножение и деление оказываются чуть экономичнее сложения); экономичным (сравнимым с затратами на одно сложение) является взятие квадратного корня;
- затраты на обращение к генератору *RAND*, используемому в системе NMPUD (метод вычетов с множителем $Q = 5^{17}$ и контролируемой мантиссой длины $m = 40$), примерно в 5–6 раз больше затрат на одно сложение (умножение, деление, взятие квадратного корня) и примерно равны затратам на вычисление функций синус и косинус;
- «умеренно большими» (вместе с затратами на обращение к генератору *RAND* и к вычислению синуса и косинуса) можно назвать затраты на вычисление обратных тригонометрических функций и логарифмов – примерно в 7–10 раз больше затрат на одно сложение (умножение, деление, взятие квадратного корня);
- «экстремально большими» являются затраты на вычисление степенной и показательной функций – в десятки (от 30 до 80) раз больше затрат на одно сложение (умножение, деление, взятие квадратного корня); использование этих функций в компьютерных расчетах желательно ограничивать.

Полученные результаты во многом объясняют повышенную, неприемлемую для практических расчетов трудоемкость многих формул метода обратной функции распределения $\xi_0 = F_{\xi}^{-1}(\alpha_0)$; $\alpha_0 \in U(0,1)$.

4. Двусторонний метод исключения как альтернатива методу обратной функции распределения.

Для трудоемких формул метода обратной функции распределения $\xi_0 = F_{\xi}^{-1}(\alpha_0)$ в научной группе А. В. Войтишека рассмотрена универсальная (во всяком случае, для конечных интервалов распределения (a, b)) технология двустороннего метода исключения с кусочно-постоянными мажорантой и минорантой и с уравниванием вероятностей при моделировании вспомогательной случайной величины методом дискретной суперпозиции и создана компьютерная система EDSRM (Economical Double-Sided Rejection Method).

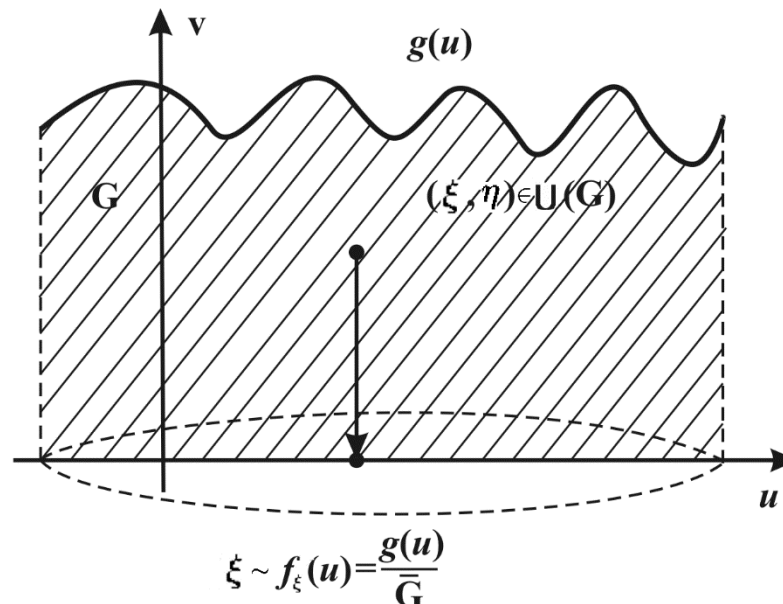
ПРЕДСТАВЛЕНИЯ ДВУСТОРОННЕГО АЛГОРИТМА И СИСТЕМЫ EDSRM ЯВЛЯЮТСЯ ПРЕДМЕТОМ ДАННОГО ДОКЛАДА.

5. Обоснование мажорантного метода исключения. Двусторонний метод.

Пусть требуется построить алгоритм численного моделирования выборочного значения ξ_0 случайной величины ξ , распределенной на отрезке (конечном интервале) в области $[a, b]$ согласно плотности $f_\xi(u)$, которая пропорциональна заданной неотрицательной функции $g(u)$, т. е.

$$f_\xi(u) = \frac{g(u)}{\bar{G}}, \quad u \in [a, b], \quad \bar{G} = \int_a^b g(u) du. \quad (*)$$

УТВЕРЖДЕНИЕ 1. Пусть случайная точка (ξ, η) равномерно распределена в области $G = \{(u, v): u \in [a, b]; 0 < v < g(u)\}$, т. е. в подграфике функции $g(u)$ из соотношения $(*)$ (обозначение $(\xi, \eta) \in U(G)$; при этом $\xi \in [a, b]$ и $\eta \in (0, g(\xi))$). Тогда случайная величина ξ распределена согласно плотности $(*)$.

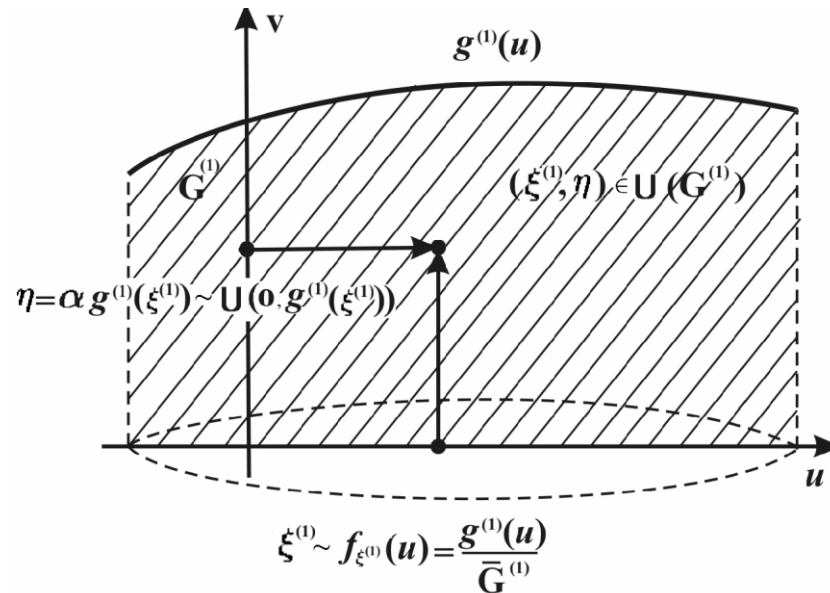


УТВЕРЖДЕНИЕ 2. Пусть случайный вектор $\xi^{(1)}$ распределен согласно плотности

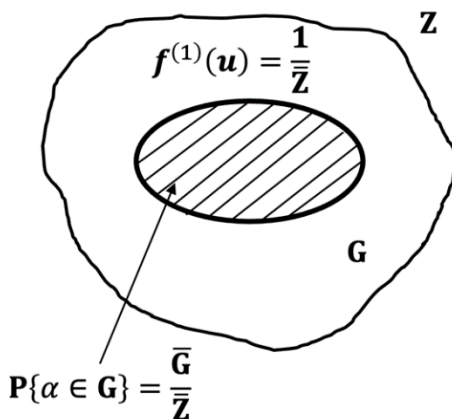
$$f_{\xi^{(1)}}(u) = \frac{g^{(1)}(u)}{\bar{G}^{(1)}}, \bar{G}^{(1)} = \int_X g^{(1)}(u) du,$$

а условное распределение при фиксированном значении $\xi^{(1)} = \xi_0^{(1)}$ случайной величины η является равномерным в интервале $\left(0, g^{(1)}(\xi_0^{(1)})\right)$.

Тогда случайная точка $(\xi^{(1)}, \eta)$ равномерно распределена в «подграфике» $G^{(1)} = \{u \in [a, b], 0 < v < g^{(1)}(u)\}$ функции $g^{(1)}(u)$, т. е. $(\xi^{(1)}, \eta) \in U(G^{(1)})$.

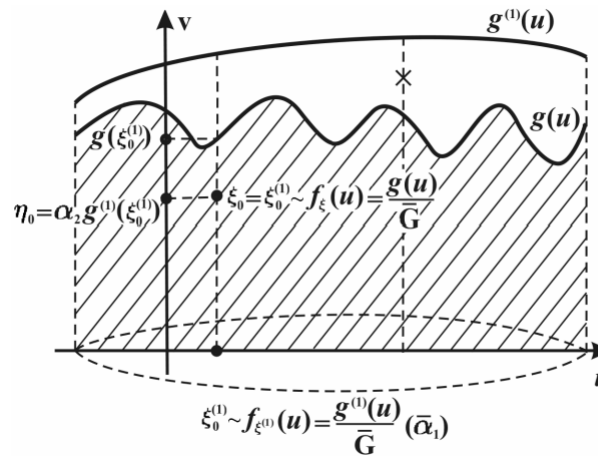


УТВЕРЖДЕНИЕ 3. Если d -мерная случайная точка $\alpha \in U(Z)$ равномерно распределена в области $Z \in R^d$ конечного объема (площади) $\bar{Z} = \int_Z du$, то она также равномерно распределена в любой подобласти $G \in Z$ объема $\bar{G}$ при условии попадания в эту подобласть; при этом $P\{\alpha \in G\} = \bar{G}/\bar{Z}$.



Для «существенной части» $g(u)$ плотности $f_{\xi}(u)$ (для которой строится алгоритм моделирования) строим мажоранту $g^{(1)}(u): g(u) \leq g^{(1)}(u)$.

Главное требование к мажоранте $g^{(1)}(u)$ таково, что для плотности $f_{\xi}(u)$ имеется эффективный алгоритм (формула) вида $\xi_0^{(1)} = \Psi^{(1)}(\bar{\alpha}_1)$ для моделирования выборочного значения случайного вектора $\xi^{(1)}$; здесь $\bar{\alpha}_1 = (\alpha_{1,1}, \dots, \alpha_{1,K})$; $\alpha_{1,j} \in U(0,1)$; $j = 1, \dots, K$ – набор стандартных случайных чисел (компьютерных реализаций – с помощью соответствующих генераторов – случайной величины $\alpha \in U(0,1)$, равномерно распределенных в интервале $(0,1)$). Это дает **мажорантный метод исключения**.



АЛГОРИТМ 1. 1. Моделируем выборочное значение $\xi_0^{(1)}$ случайного вектора (случайной величины) $\xi^{(1)}$ согласно плотности $f_{\xi^{(1)}}(u)$: $\xi_0^{(1)} = \Psi^{(1)}(\bar{\alpha}_1)$, а также значение $\eta_0 = \alpha_2 g^{(1)}(\xi_0^{(1)})$; $\alpha_2 \in U(0,1)$.

Согласно утверждению 2, $(\xi_0^{(1)}, \eta_0) \in U(\xi_0^{(1)})$.

2. Если $\eta_0 < g(\xi_0^{(1)})$, то, согласно утверждению 3, $(\xi_0^{(1)}, \eta_0) \in U(G)$, и, согласно утверждению 1, величину $\xi_0^{(1)}$ можно принять в качестве искомого выборочного значения, имеющего нужную плотность распределения $f_{\xi}(u)$:

$$\xi_0 = \xi_0^{(1)}.$$

В случае, когда неравенство $\eta_0 < g(\xi_0^{(1)})$ не выполнено, повторяем пункт 1 данного алгоритма и т. д.

Следующая модификация алгоритма 1 эффективна в достаточно распространенном случае, когда требуется моделировать выборочное значение случайной величины ξ , плотность которой пропорциональна функции $g(u)$, **вычисление значений которой весьма трудоемко**. В этом случае помимо мажоранты $g^{(1)}(u)$ строим миноранту $g^{(2)}(u)$, такую, что $0 \leq g^{(2)}(u) \leq g(u) \leq g^{(1)}(u); u \in [a, b]$.

АЛГОРИТМ 2. 1. Моделируем выборочное значение $\xi_0^{(1)} = \Psi^{(1)}(\bar{\alpha}_1)$ согласно плотности $f_{\xi^{(1)}}(u)$, а также значение $\eta_0 = \alpha_2 g^{(1)}(\xi_0^{(1)})$.

2. Вместо неравенства $\eta_0 < g(\xi_0^{(1)})$ проверяем сначала соотношение

$$\eta_0 \leq g^{(2)}(\xi_0^{(1)}).$$

Если оно выполнено, то пара $(\xi_0^{(1)}, \eta_0)$ принадлежит «подграфику» функции $g^{(2)}(u)$, а значит, и области G . Тогда можно положить $\xi_0 = \xi_0^{(1)}$.

В случае же $\eta_0 > g^{(2)}(\xi_0^{(1)})$ проверяем неравенство $\eta_0 < g(\xi_0^{(1)})$. Если оно выполнено, то $\xi_0 = \xi_0^{(1)}$, иначе повторяется пункт 1 данного алгоритма и т.д.

В связи с соотношением $g^{(2)}(u) \leq g(u) \leq g^{(1)}(u)$ алгоритм 2 называют **двусторонним методом исключения**.

В качестве функций $g^{(2)}(u)$ и $g^{(1)}(u)$ в данной работе используются **кусочно-постоянные приближения снизу и сверху** для функции $g(u)$.

6. Использование кусочно-постоянных мажорант и минорант (для монотонных плотностей на конечных интервалах распределения).

При выборе функций $g^{(1)}(u), g^{(2)}(u)$ таких, что

$$0 \leq g^{(2)}(u) \leq g(u) \leq g^{(1)}(u),$$

наиболее универсальным и целесообразным является использование кусочно-постоянных приближений функции $g(u)$ снизу и сверху.

Для монотонно возрастающей функции (например, для функции, $g(u) = u^s, 0 \leq u \leq 1$, соответствующей плотности степенного распределения) можно взять

$$g^{(1)}(u) \equiv g(u_i) = A_i^{(+)} (= u_i^s); g^{(2)}(u) \equiv g(u_{i-1}) = A_i^{(-)} (= u_{i-1}^s); \\ u \in \Delta_i \in (u_{i-1}, u_i]; i = 1, \dots, M,$$

для сетки

$$a = u_0 < u_1 < \dots < u_{M-1} < u_M = b.$$

7. Моделирование вспомогательной случайной величины по кусочно-постоянной плотности, пропорциональной мажоранте. Использование теории модифицированного метода дискретной суперпозиции и уравнивания вероятностей.

Для построенной кусочно-постоянной мажоранты $g^{(1)}(u)$ плотность $f_{\xi^{(1)}}(u)$ случайной величины $\xi^{(1)} \in [a, b]$ можно представить в виде составной плотности

$$f_{\xi^{(1)}}(u) = \sum_{i=1}^M p_i g_i(u) \chi^{\Delta_i}(u), \text{ где } p_i = \frac{A_i^{(+)}(u_i - u_{i-1})}{\sum_{j=1}^M [A_j^{(+)}(u_j - u_{j-1})]}; \quad g_i(u) \equiv \frac{1}{u_i - u_{i-1}}$$

и $\chi^{\Delta_i}(u)$ – индикатор полуинтервала Δ_i , т. е. $\chi^{\Delta_i}(u) \equiv 1$ при $u \in \Delta_i$ и $\chi^{\Delta_i}(u) = 0$ иначе.

Далее можно использовать соответствующую версию алгоритма метода суперпозиции.

АЛГОРИТМ 3. 1. Реализуя стандартное случайное число $\alpha_0 \in U(0,1)$ и используя наиболее экономичный из алгоритмов моделирования целочисленной случайной величины μ с распределением вида $\mathbf{P}\{\mu = i\} = p_i; i = 1, \dots, M$ для вероятностей $\{p_i\}$, получаем значение $\mu_0 = m$.

2. Моделируем выборочное значение $\xi_0^{(1)}$ случайной величины $\xi^{(1)}$ согласно плотности равномерного распределения $g_m(u)$ по формуле

$$\xi_0^{(1)} = u_{m-1} + (u_m - u_{m-1})\alpha_2; \quad \alpha_2 \in U(0,1).$$

Выберем сетку $-\infty < a = u_0 < u_1 < \dots < u_{M-1} < u_M = b < +\infty$ таким образом, что площади $S_i = A_i^{(+)}(u_i - u_{i-1}) = g(u_i)(u_i - u_{i-1}) \equiv S$ одинаковы для $i = 1, \dots, M$.

Такая сетка может быть построена с помощью процедуры «подбора» (с компьютерной точностью), соответствующей следующей процедуре.

Определим процедуру-функцию от $a_M = r$ (самого правого из левых концов полуинтервалов $(u_{i-1}, u_i]$) – обозначим ее $z(r)$ – как следующую последовательность команд в коде Maple:

$z^{(rej)}(r) : u_{M-1} = r; S^{(rej)} = (b - r)g(b); \text{for } i \text{ from } M - 2 \text{ by } -1 \text{ to } 1 \text{ do } u_i =$
 $u_{i+1} - \frac{S^{(rej)}}{g(u_{i+1})} \text{ od ; return } (S^{(rej)} - (u_1 - a)g(u_1))$

«Двигаем» (подбираем) параметр r пока с высокой точностью не будем иметь $z^{(rej)}(r) \approx 0$.

АЛГОРИТМ 4. Выберем $\Delta = \frac{b-a}{M}$ и возьмем $r_0 = b - \Delta$ и вычислим $z^{(rej)}(r_0)$. Для монотонно возрастающей функции имеем $z^{(rej)}(r_0) > 0$, т. е. площадь S «излишне большая» и параметр $r = r_0$ нужно увеличить.

Пусть стандартное представление числа $\Delta = a \times 10^k; 1 \leq a \leq 10$ имеет порядок k (это целое – положительное или отрицательное – число). Возьмем $\varepsilon^{(1)} = 10^k$ и начнем прибавлять эту величину к r_0 до тех пор пока первый раз не получим точку r_1 такую, что $z^{(rej)}(r_1) < 0$. Теперь возьмем $\varepsilon^{(2)} = 10^{k-1}$ и начнем вычитать эту величину из r_1 до тех пор пока первый раз не получим точку r_2 такую, что $z^{(rej)}(r_2) > 0$ и т. д.

Процедуру продолжаем вплоть до номера итерации s такого, что $\varepsilon^{(s)} < 10^{-L}$ (например, для $L = 16$), и полагаем $r = r_s$.

Наконец, можно использовать следующий факт.

УТВЕРЖДЕНИЕ 4. Если в алгоритме 3 получено $\mu_0 = m$ (т. е. $\alpha \in [R_{m-1}, R_m)$), то случайная величина $\beta(\alpha) = \frac{\alpha - R_{m-1}}{p_m}$ равномерно распределена в интервале $(0,1)$ (т.е. $\beta(\alpha) \in U(0,1)$); при этом величины μ и $\beta(\alpha)$ независимы.

Из утверждения 4 следует, что вместо алгоритма 3 можно применить т. н. **модифицированный метод дискретной суперпозиции**, означающий замену величины α_2 из второго пункта алгоритма 3 на величину $\frac{\alpha_1 - R_{m-1}}{p_m}$. Отметим также, что выполнении соотношения $S_i \equiv S$ получается

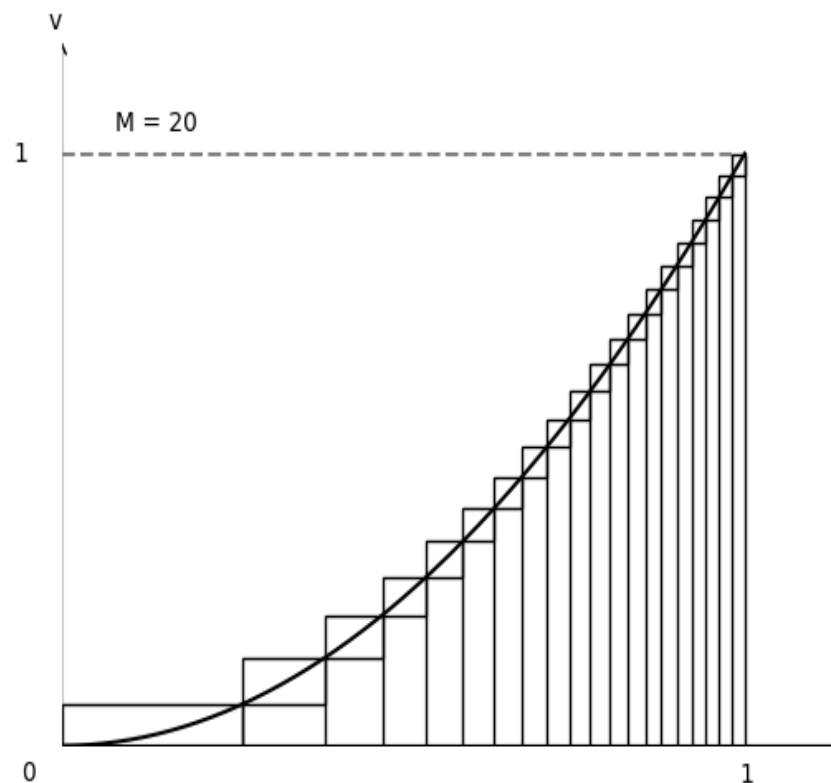
$$R_i = \frac{i}{M}.$$

Для такого варианта моделирование значения $\xi_0^{(1)}$ из первого пункта алгоритма метода исключения может быть реализовано по **рекордно экономичным формулам**

$$m = [M\alpha_1] + 1;$$

$$\xi_0^{(1)} = u_{m-1} + (u_m - u_{m-1})(M\alpha_1 - m + 1)$$

для $\alpha_1 \in U(0, 1)$.



АЛГОРИТМ 5. Предварительно формируем массив $U = \{u_0, u_1, \dots, u_{M-1}, u_M\}$, а также массивы минимумов и максимумов функции $g(u)$ на элементах Δ_i :

$$\mathbf{MIN} = \{A_1^{(-)} = g(u_0), A_2^{(-)} = g(u_1), \dots, A_M^{(-)} = g(u_{M-1})\} \text{ и}$$

$$\mathbf{MAX} = \{A_1^{(+)} = g(u_1), A_2^{(+)} = g(u_2), \dots, A_M^{(+)} = g(u_M)\}.$$

1. Моделируем выборочное значение $\xi_0^{(1)}$ согласно формулам $m = [M\alpha_1] + 1$, $\xi_0^{(1)} = u_{m-1} + (u_m - u_{m-1})(M\alpha_1 - m + 1)$, а также значение $\eta_0 = \alpha_2 A_m^{(+)} \in U(0, A_m^{(+)})$; здесь $\alpha_2 \in U(0, 1)$. Согласно утверждению 2, точка $(\xi_0^{(1)}, \eta_0)$ равномерно распределена в подграфике $G^{(1)} = \{(u, v) : u \in [a, b]; 0 < v < g^{(1)}(u)\}$.

2. Сравниваем сначала значение η_0 с соответствующим значением кусочно-постоянной миноранты $g^{(2)}(u)$, т. е. проверяем неравенство $\eta_0 < A_m^{(-)}$. Если оно выполнено, то пара $(\xi_0^{(1)}, \eta_0)$ принадлежит подграфику функции $g^{(2)}(u)$, а значит, и области $G = \{(u, v) : u \in [a, b]; 0 < v < g(u)\}$. Тогда, согласно утверждениям 1 и 3, можно положить $\xi_0 = \xi_0^{(1)}$.

3. Если же $\eta_0 \geq A_m^{(-)}$, то проверяем неравенство $\eta_0 < g(\xi_0^{(1)})$. Если оно выполнено, то $(\xi_0^{(1)}, \eta_0) \in G$, и тогда, согласно утверждениям 1 и 3, можно положить $\xi_0 = \xi_0^{(1)}$. В противном случае (т. е. при $\eta_0 \geq g(\xi_0^{(1)})$) повторяем пункт 1 и т. д. — до получения значения ξ_0 .

Нами также разработана компьютерная система EDSRM (Economical Double-Sided Rejection Method) – ссылка <http://edsrm.nikita-e.ru/>. Система содержит **библиотеку основных разработанных алгоритмов и удобную диалоговую систему.**

Для исследуемой плотности $f_{\xi}(u) = Hg(u); -\infty < a \leq u \leq b < +\infty; H = \frac{1}{\int_a^b g(v) dv} \quad (*)$

случайной величины $\xi \in [a, b]$ диалоговая система EDRSM позволяет:

- вносить в систему формулу функции $g(u)$ и границы отрезка $[a, b]$;
- вносить формулы метода обратной функции распределения для плотности $(*)$;
- проводить сравнение (с построением соответствующих диаграмм) затрат двустороннего алгоритма метода исключения и формулы обратной функции распределения для различных монотонных функций $g(u)$ для различных значений параметров n (это количество моделируемых выборочных значений случайной величины $\xi \in [a, b]$) и M (это количество интервалов разбиения для неравномерной сетки $U = \{u_0 = a, u_1, u_2, \dots, u_{M-1}, u_M = b\}$, используемой при уравнивании вероятностей); к слову, рекомендуемыми нами являются значения $n = 10^8$ (это значение дает показательное значение затрат используемых алгоритмов и формул) и $M = 330$ (для широкого класса распределений для $M > 330$ не происходило сколько-нибудь существенного увеличения затрат двустороннего алгоритма – это и определяет его универсальность и экономичность);
- наблюдать результаты построения кусочно-постоянных мажорант и минорант на неравномерных сетках $U = \{u_0 = a, u_1, u_2, \dots, u_{M-1}, u_M = b\}$, полученных после процедуры уравнивания вероятностей, для различных M на соответствующих графиках;
- наблюдать формирование гистограммы для моделируемого распределения и ее приближение (с ростом числа n) к исследуемой монотонной плотности распределения и др.

ЗАКЛЮЧЕНИЕ

Построен универсальный (во всяком случае, для случайных величин, распределенных на ограниченных интервалах – отрезках – согласно непрерывным кусочно-монотонным плотностям) экономичный алгоритм двустороннего метода исключения с кусочно-постоянными мажорантой и минорантой.

Предложена технология уравнивания вероятностей, основанная на построении специальных неравномерных сеток для формирования необходимых кусочно-постоянных мажорант. Эта технология определяет рекордную экономичность и универсальность предложенного двустороннего алгоритма метода исключения.

Разработана компьютерная система EDSRM, предоставляющая пользователю коды нового алгоритма на широко распространенном языке программирования С и позволяющая сравнивать – в удобном диалоговом формате – затраты построенного алгоритма с трудоемкостями соответствующих формул метода обратной функции распределения.

Имеется публикация

Брызгалов В. Л., Войтишек А. В. Анализ трудоемкости формул метода обратной функции распределения для случайных величин с конечным интервалом распределения // Информационные технологии и математическое моделирование (ИТММ-2023): Материалы XXII Международной конференции им. А. Ф. Терпугова (4–9 декабря 2023 года). – Томск: Издательство Томского государственного университета, 2023. – Часть 1. – С. 314 – 321

+ готовятся более «основательные» и общедоступные статьи.

БЛИЖАЙШИЕ ПЛАНЫ ИССЛЕДОВАНИЙ

1. Детально изучить построенный двусторонний алгоритм метода исключения для случайной величины, имеющей **монотонную плотность на отрезке** (основной пример – **степенное распределение** $f_{\xi}(u) = (s + 1)u^s; 0 \leq u \leq 1, s > 0$).

2. Провести аналогичное исследование для случайной величины, имеющей **кусочно-монотонную плотность на отрезке** (основной пример – **бета-распределение** $f_{\xi}(u) = Hu^{\mu-1}(1-u)^{\nu-1}; 0 \leq u \leq 1, \mu > 1, \nu > 1$).

3. Провести критический анализ и модернизацию (в сторону корректности использования) т. н. **зигкурат-метода** – см.

Marsaglia G., Tsang W. W. The ziggurat method for generating random variables. Journal of Statistical Software. 2000. 5 (8)

(это «положенный набор» двусторонний метод, дающий возможность избавиться от одного обращения в датчику *RAND*) – **ПРЕДМЕТ ДОКЛАДА ЧЕРЕЗ ГОД (?!)**.

4. Провести анализ применения алгоритмов с уравниванием вероятностей (двустороннего алгоритма и зигкурат-метода) для **практически значимого экспоненциального распределения** $f_{\xi}(u) = \lambda e^{-\lambda u}; u > 0, \lambda > 0$ (это первый пример **распределения «с хвостом»**, определяющим специфику конструкций алгоритмов).

5. Провести анализ применения алгоритмов с уравниванием вероятностей (двустороннего алгоритма и зигкурат-метода) для других **важных распределений «с хвостами»**; в частности, для **гауссовского распределения** $f_{\xi}(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$ и для **гамма-распределения** $f_{\xi}(u) = Hu^{\nu-1}e^{-\lambda u}; u > 0, \nu > 0, \lambda > 0$.

СПАСИБО ЗА ВНИМАНИЕ!

vav@osmf.sccc.ru